GHAJAR
EXHIBIT 42

Page 1

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN FRANCISCO DIVISION

_____

Richard Kadrey, et al.,   )

Individual and Representative)

Plaintiffs,   )   CASE NO.

)   3:23-cv-03417-VC

          -against-   )

)

Meta Platforms, Inc.,   )

     Defendant.   )

_____)

*** HIGHLY CONFIDENTIAL ***

ATTORNEYS' EYES ONLY

VIDEO-RECORDED DEPOSITION OF

MELANIE KAMBADUR

Cooley, LLP

55 Hudson Yards

New York, New York   10001

09/17/2024

9:07 a.m. (EDT)

REPORTED BY: MONIQUE CABRERA

_____

DIGITAL EVIDENCE GROUP

1730 M Street, NW, Suite 812

Washington, D.C. 20036

(202) 232-0646

Page 11

```
 1              A.     Yes.  My full name is

 2       Melanie Rae Kambadur.

 3              Q.     And your address was

[REDACTED]

 5              A.     Yes.

 6              Q.     Okay.  And where are you presently

 7       employed?

 8              A.     At Meta.

 9              Q.     Meta Platforms?

10              A.     Yes.

11              Q.     Okay.  And what is your current

12       position at Meta?

13              A.     I'm a research engineering manager.

14              Q.     Is there a particular division you

15       work for?

16              A.     I work in the gen AI org.

17                     (Reporter clarification.)

18                     THE WITNESS:  Gen AI.

19       BY MR. YOUNG:

20              Q.     And how long have you been in that

21       role?

22              A.     I have been a research manager at
```

9/17/2024   Richard Kadrey, et al. v. Meta Platforms, Inc.   Melanie Kambadur
Highly Confidential - Attorneys' Eyes Only

Page 12

```
 1    Meta for three years in -- since March.

 2    Subtracting six.  Six months.  And in the gen AI

 3    organization for about the last year and a half.

 4         Q.    I'm sorry.  You said you were in the

 5    gen AI division for the last year and a half?

 6         A.    Yes.  Approximately.

 7         Q.    And what division were you in before

 8    you were in the gen AI division?

 9         A.    In a group called FAIR.

10         Q.    And that's Facebook AI Research?

11         A.    Yes.  I believe they changed it to

12    Fundamental AI Research at some point as well.

13         Q.    But when you were there, it was

14    Facebook AI Research?

15         A.    I don't recall exactly when they

16    flipped the name.

17         Q.    Okay.  Have you ever had your

18    deposition taken before?

19         A.    No.

20         Q.    All right.  And you may have covered

21    some of this with your counsel in preparation for

22    today.  I want to -- but I want to just go over
```

Page 70

```
 1      We said he worked on pre-processing the data,

 2      that is fair to say.

 3           Q.    What is pre-process -- processing?

 4      Excuse me.

 5           A.    That means transforming the data

 6      with a series of steps that make it more suitable

 7      for training models.

 8           Q.    Does that include removing -- strike

 9      that.

10                 What kind of steps would be involved

11      in pre-processing?

12           A.    Are you asking specifically for

13      Llama models or more generally?

14           Q.    Yeah, for the whole Llama models.

15           A.    I don't know all the steps.  But for

16      example, we perform various types of

17      de-duplication.  We tokenize the data, which

18      means -- yeah, we tokenize the data.  We chunk

19      the data into spans of text.

20           Q.    Let's talk about de-duplication.  So

21      if the models -- if a model is trained or

22      pre-trained on the same text over and over again,
```

Page 71

1   would that affect the model's performance?

2   A.   It could.

3   Q.   How could it affect the model's

4   performance?

5   A.   It could make the training less

6   efficient to repeatedly seem -- see the same

7   data.

8   Q.   So you want to limit the model's

9   exposure to common text?

10   Would that be fair to say?

11   A.   What do you mean by "common text"?

12   Q.   For example -- so de-duplication,

13   right, you would want to remove text that appears

14   frequently from the pre-training dataset.

15   Is that -- would that be fair to

16   say?

17   A.   It depends on the text.

18   Q.   So, for example, if you -- you know,

19   you -- you've read books, correct?

20   A.   Yes, I've read books.

21   Q.   Okay.   So you know in the very front

22   of the page, there is the page including all the

Page 72

1   copyright information? For example, "published

2   by," all the have information.

3   A.   Yes.

4   Q.   Would that be the type of

5   information you would remove before processing?

6   A.   Potentially.

7   Q.   Was that information actually

8   removed for data used to pre-train the Llama

9   models?

10   A.   I don't recall.

11   Q.   Do you know -- going back to Libgen.

12   Do you know if it contained fiction

13   books?

14   A.   I believe it did.

15   Q.   How about non-fiction books?

16   A.   Yes, I believe it did.

17   Q.   Are you familiar with the Project

18   Gutenberg?

19   A.   At a very high level, yes.

20   Q.   What do you understand Project

21   Gutenberg to be?

22   A.   A -- I -- I don't know the full

9/17/2024   Richard Kadrey, et al. v. Meta Platforms, Inc.   Melanie Kambadur
Highly Confidential - Attorneys' Eyes Only

Page 337

```
 1    CERTIFICATE OF SHORTHAND REPORTER NOTARY PUBLIC

 2              I, Monique Cabrera, the officer

 3        before whom the foregoing deposition was

 4        taken, do hereby certify that the foregoing

 5        transcript is a true and correct record of

 6        the testimony given; that said testimony was

 7        taken by me stenographically and thereafter

 8        reduced to typewriting under my direction;

 9        and that I am neither counsel for, related

10        to, nor employed by any of the parties to

11        this case and have no interest, financial or

12        otherwise, in its outcome.

13              IN WITNESS WHEREOF, I have hereunto

14        set my hand this 17th day of September, 2024.

15

16

17

18        _____

19        MONIQUE CABRERA

20        Notary Public in and for the State of New York

21        County of Suffolk

          My Commission No.

22        Expires:  06/12/2026
```